

A Variational Perspective on Diffusion-Based Generative Models and Score Matching

Year: 2021 | Citations: 226 | Authors: Chin-Wei Huang, Jae Hyun Lim, Aaron C. Courville

Abstract

Discrete-time diffusion-based generative models and score matching methods have shown promising results in modeling high-dimensional image data. Recently, Song et al. (2021) show that diffusion processes that transform data into noise can be reversed via learning the score function, i.e. the gradient of the log-density of the perturbed data. They propose to plug the learned score function into an inverse formula to define a generative diffusion process. Despite the empirical success, a theoretical underpinning of this procedure is still lacking. In this work, we approach the (continuous-time) generative diffusion directly and derive a variational framework for likelihood estimation, which includes continuous-time normalizing flows as a special case, and can be seen as an infinitely deep variational autoencoder. Under this framework, we show that minimizing the score-matching loss is equivalent to maximizing a lower bound of the likelihood of the plug-in reverse SDE proposed by Song et al. (2021), bridging the theoretical gap.